

A Difference Scheme for the Numerical Solution of an Advection Equation with Aftereffect

S. I. Solodushkin^{1*}
(Submitted by V.V. Vasin)

¹Ural Federal University, ul. Turgeneva 4, Ekaterinburg, 620000 Russia

Received March 23, 2013

Abstract—We propose a family of grid methods for the numerical solution of an advection equation with a time delay in a general form. The methods are based on the idea of separating the current state and the prehistory function. We prove the convergence of the second-order method coordinatewise and do that of the first-order with respect to time. The proof is based on techniques applied for proving analogous theorems for functional differential equations and on the general theory of difference schemes. We illustrate the obtained results with a test example.

DOI: 10.3103/S1066369X13100095

Keywords and phrases: *advection equations, time delay, difference scheme, numerical methods.*

1. Introduction. The problem. Many papers are dedicated to the qualitative theory of partial functional differential equations (see, for example, [1] and references therein). Since in most cases one cannot solve such equations analytically, the elaboration, substantiation, and computer realization of numerical methods for this class of equations are of essential interest (see the review in [2]). Many difference schemes [3] are known for advection equations without delay. In the paper [4], for an advection equation with a delay, one considers an approximation of the derivative with respect to the phase variable with two nodes, which provides only the first order of convergence with respect to x . The present paper continues the investigation initiated in [5].

Advection equations with delay arise in modeling the dynamics of populations structured with respect to the cell size [6, 7], the age of specimen, etc.

Consider the following advection equation with aftereffect:

$$\frac{\partial u(x, t)}{\partial t} + a \frac{\partial u(x, t)}{\partial x} = f(x, t, u(x, t), u_t(x, \cdot)), \quad (1)$$

where $x \in [0, X]$ and $t \in [t_0; \theta]$ are, respectively, the spatial and time variables (independent ones), $u(x, t)$ is the desired function, $u_t(x, \cdot) = \{u(x, t + \xi), -\tau \leq \xi < 0\}$ is the prehistory function of the desired function at the moment t , $\tau > 0$ is the value of the delay, and $a > 0$ is a coefficient.

Together with the equation, we state the initial condition

$$u(x, t) = \varphi(x, t), \quad x \in [0, X], \quad t \in [t_0 - \tau, t_0], \quad (2)$$

the boundary one

$$u(0, t) = g(t), \quad t \in [t_0, \theta], \quad (3)$$

and the fitting condition

$$g(t_0) = \varphi(0, t_0).$$

We assume that the functional f and functions φ and g are such that problem (1)–(3) has a unique solution. Questions of the existence and uniqueness of a solution to the stated boundary value problem were considered in [1].

*E-mail: solodushkin_s@mail.ru.

We denote by $Q = Q[-\tau, 0)$ the set of piecewise-continuous on $[-\tau, 0)$ functions $u(\xi)$ which have a finite number of discontinuity points of the first kind, where a function is continuous from the right. We define the norm of functions on Q by the formula $\|u(\xi)\| = \sup_{\xi \in [-\tau, 0)} |u(\xi)|$. Additionally we assume that the functional $f(x, t, u, v(\cdot))$ is defined on $[0, X] \times [t_0, \theta] \times R \times Q$ and is Lipschitzian in the last two arguments, i.e.,

$$\exists L_f \in R \quad \forall x \in [0, X], \quad t \in [t_0, \theta], \quad u^1 \in R, \quad u^2 \in R, \quad v^1 \in Q, \quad v^2 \in Q : \\ |f(x, t, u^1, v^1(\cdot)) - f(x, t, u^2, v^2(\cdot))| \leq L_f(|u^1 - u^2| + \|v^1(\cdot) - v^2(\cdot)\|_Q).$$

2. The difference scheme. We split the segment of variation of the spatial variable $[0, X]$ into parts with the step $h = X/N$, introducing points $x_i = ih$, $i = 0, \dots, N$, and do the segment of variation of the time variable $[t_0, \theta]$ into parts with the step Δ , introducing points $t_j = t_0 + j\Delta$, $j = 0, \dots, M$. We assume that the value $\tau/\Delta = m$ is a natural number. We denote by u_j^i approximations of the function $u(x_i, t_j)$ at nodes.

For the calculation of the functional f it is necessary to know values of the function $u(x, t)$ between the grid nodes. With each fixed $i = 0, \dots, N$ we introduce the discrete prehistory up to the time moment t_j , $j = 0, \dots, M$: $\{u_k^i\}_j = \{u_k^i, j - m \leq k \leq j\}$. We treat the operator defined on the set of all feasible prehistories and acting by the rule $I : \{u_k^i\} \rightarrow v^{i,j}(\cdot) \in Q[-\tau, \Delta)$ as the interpolation-extrapolation operator. Here $Q[-\tau, \Delta)$ is the set of piecewise continuous on $[-\tau, \Delta)$ functions with a finite number of discontinuity points of the first kind, where a function is continuous from the right.

We say that an interpolation-extrapolation operator has an error of order p on the exact solution, if

$$\exists C_1 \in R, \quad C_2 \in R \quad \forall i = 0, \dots, N, \quad j = 0, \dots, M, \quad t \in [t_j - \tau, t_{j+1}] : \\ |v^{i,j}(t) - u(x_i, t)| \leq C_1 \max_{j-m \leq k \leq j} |u_k^i - u(x_i, t)| + C_2 \Delta^p.$$

For example, a piecewise-constant interpolation has the first order, and a piecewise-linear one

$$v^{i,j}(t_j + \xi) = \frac{1}{\Delta}((t_k - t_j - \xi)u_{k-1}^i + (t_j + \xi - t_{k-1})u_k^i), \quad t_{k-1} \leq t_j + \xi \leq t_k, \quad -\tau \leq \xi \leq 0,$$

with an extrapolation by continuity

$$v^{i,j}(t_j + \xi) = \frac{1}{\Delta}((-\xi)u_{j-1}^i + (\Delta + \xi)u_j^i), \quad 0 \leq \xi \leq \Delta,$$

has the second order ([8], P. 98).

For $0 \leq s \leq 1$ we consider the family of methods $j = 0, \dots, M - 1$:

$$\frac{u_{j+1}^1 - u_j^1}{\Delta} + a \left(s \frac{-4u_{j+1}^0 - \frac{2h}{a}(f_{j+1}^0 - \dot{g}_{j+1}) + 4u_{j+1}^1}{2h} + (1-s) \frac{-4u_j^0 - \frac{2h}{a}(f_j^0 - \dot{g}_j) + 4u_j^1}{2h} \right) = f_j^1, \\ \frac{u_{j+1}^i - u_j^i}{\Delta} + a \left(s \frac{u_{j+1}^{i-2} - 4u_{j+1}^{i-1} + 3u_{j+1}^i}{2h} + (1-s) \frac{u_j^{i-2} - 4u_j^{i-1} + 3u_j^i}{2h} \right) = f_j^i, \quad i = 2, \dots, N, \quad (4)$$

with the initial condition

$$u_0^i = \varphi(x_i, t_0), \quad i = 0, \dots, N, \\ v^{i,0}(t) = \varphi(x_i, t), \quad t < t_0, \quad i = 0, \dots, N,$$

and the boundary one

$$u_j^0 = g_0(t_j), \quad j = 0, \dots, M.$$

Here $f_j^i = f(x_i, t_j, u_j^i, v^{i,j}(\cdot))$ is the value of the function f calculated on an approximate solution, $v^{i,j}(\cdot)$ is the result of a piecewise-linear interpolation with an extrapolation by continuity, and $\dot{g}_j =$

$\frac{dg(t)}{dt}\big|_{t=t_0+j\Delta}$. For constructing a numerical method, we additionally assume that $g(t)$ is a differentiable function.

Let us explain the way in which we have obtained the scheme. The derivative $\partial u/\partial t$ in Eq. (1) is approximated by a finite difference over two nodes. For nodes (i, j) , $i = 2, \dots, N$, $j = 0, \dots, M-1$, the derivative $\partial u/\partial x$ is approximated by a finite difference over three nodes on the right edge. For $i = 1$ such an approximation requires to calculate u_j^{-1} . For $i = 1$ we apply the approximation over three nodes with the multiple node $(0, j)$

$$\frac{\partial u_j^1}{\partial x} \approx \frac{-4u_j^0 - 2h \partial u_j^0/\partial x + 4u_j^1}{2h}.$$

In view of (1) we have $\frac{\partial u_j^0}{\partial x} = \frac{1}{a}(f_j^0 - \frac{\partial u_j^0}{\partial t})$; due to (3) we obtain $\frac{\partial u_j^0}{\partial x} = \frac{1}{a}(f_j^0 - \dot{g}_j)$.

We call the mesh function

$$\begin{aligned} \Psi_j^1 &= \frac{u(x_1, t_{j+1}) - u(x_1, t_j)}{\Delta} + as \frac{-4u(x_0, t_{j+1}) - \frac{2h}{a}(f_{j+1}^0 - \dot{g}_{j+1}) + 4u(x_1, t_{j+1})}{2h} \\ &\quad + a(1-s) \frac{-4u(x_0, t_j) - \frac{2h}{a}(f_{j+1}^0 - \dot{g}_j) + 4u(x_1, t_j)}{2h} - \bar{f}_j^1, \\ \Psi_j^i &= \frac{u(x_i, t_{j+1}) - u(x_i, t_j)}{\Delta} + as \frac{u(x_{i-2}, t_{j+1}) - 4u(x_{i-1}, t_{j+1}) + 3u(x_i, t_{j+1})}{2h} \\ &\quad + a(1-s) \frac{u(x_{i-2}, t_j) - 4u(x_{i-1}, t_j) + 3u(x_i, t_j)}{2h} - \bar{f}_j^i, \quad i = 2, \dots, N \end{aligned} \quad (5)$$

the residue of method (4). Here $\bar{f}_j^i = f(x_i, t_j, u(x_i, t_j), u_{t_j}(x_i, \cdot))$ is the value of the function f calculated on the exact solution.

Theorem 1. Assume that the exact solution $u(x, t)$ to problem (1)–(3) is thrice continuously differentiable with respect to x and twice continuously differentiable with respect to t , while the first derivative of the solution with respect to x is continuously differentiable in t . Then the residue of method (4) has the order $h^2 + \Delta$.

Denote $\varepsilon_j^i = u(x_i, t_j) - u_j^i$, $i = 0, \dots, N$, $j = 0, \dots, M$.

Definition 1. A method is said to converge, if $\varepsilon_j^i \rightarrow 0$ as $h \rightarrow 0$ and $\Delta \rightarrow 0$ for all $i = 0, \dots, N$ and $j = 0, \dots, M$. A method is said to converge with the order of $h^p + \Delta^q$, if there exists a constant C such that $\|\varepsilon_j^i\| \leq C(h^p + \Delta^q)$ for all $i = 0, \dots, N$ and $j = 0, \dots, M$.

3. Stability and convergence order of a method. In this Item we consider problems with the homogeneous boundary condition $u(x_0, t) = 0$, $t \in [t_0, \theta]$. The replacement $\tilde{u}(x, t) = u(x, t) - g(t)$ turns the initial problem into the mentioned one.

Introduce a vector $y_j = (u_j^1, u_j^2, \dots, u_j^N)' \in Y$, $j = 0, \dots, M-1$, where the symbol $'$ is the transposition sign, and Y is the vector space with the norm

$$\|y\|^2 = \sum_{i=1}^N y_i^2.$$

On the space Y we define an operator A by the matrix

$$A = \frac{a}{2h} \begin{pmatrix} 4 & 0 & 0 & 0 & \dots & \dots & \dots & 0 \\ -4 & 3 & 0 & 0 & \dots & \dots & \dots & 0 \\ 1 & -4 & 3 & 0 & \dots & \dots & \dots & 0 \\ 0 & 1 & -4 & 3 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & 1 & -4 & 3 & 0 \\ 0 & \dots & \dots & 0 & 0 & 1 & -4 & 3 \end{pmatrix}.$$

Then we can rewrite system (4) in the form

$$\frac{y_{j+1} - y_j}{\Delta} + sAy_{j+1} + (1-s)Ay_j = F_j, \quad (6)$$

where $F_j = (f_j^1 + sf_{j+1}^0 + (1-s)f_j^0, f_j^2, \dots, f_j^N)'$.

Let us use the identity $y_{j+1} = y_j + \Delta(y_{j+1} - y_j)/\Delta$ and introduce the operator $B = E + s\Delta A$ (E is the identity operator). We can write Eq. (6) as a two-level difference scheme in the canonical form [9]

$$B \frac{y_{j+1} - y_j}{\Delta} + Ay_j = F_j. \quad (7)$$

The operator A is positive definite with eigenvalues $\lambda_1(A) = 8h/a$, $\lambda_2(A) = \dots = \lambda_n(A) = 6h/a$. Hence, B is a positive definite operator. Since B is invertible, we can write (7) in the form $y_{j+1} = Sy_j + \Delta B^{-1}F_j$, where $S = (E - \Delta B^{-1}A)$ is the so-called [8, 10] operator of moving by a step.

Definition 2. The difference scheme (7) is said to be stable, if $\|S\| < 1$.

Theorem 2. If the condition $s \geq 1/2$ is fulfilled, then the difference scheme (7) is stable.

Proof. Let us consider (7) from the point of view of operator-difference equations and apply methods of the stability verification of a two-level difference scheme [9] and the separation of finite-dimensional and infinite-dimensional components [8, 10].

We symmetrize Eq. (7) by multiplying its both parts by A^{-1} and obtain

$$(A^{-1} + s\Delta E) \frac{y_{j+1} - y_j}{\Delta} + Ey_j = A^{-1}F_j.$$

Denoting $\hat{B} = A^{-1} + s\Delta E$, $\hat{A} = E$, and $\hat{F}_j = A^{-1}F_j$, we write

$$\hat{B} \frac{y_{j+1} - y_j}{\Delta} + \hat{A}y_j = \hat{F}_j. \quad (8)$$

Method (8) is stable if and only if $\hat{B} \geq 0.5\hat{A}$ ([9], P. 333, theorem 1). This is equivalent to $A^{-1} + \Delta E(s - 0.5) \geq 0$. Since A^{-1} is a positive definite operator, the latter inequality is fulfilled for any Δ , only if $s \geq 0.5$. \square

By the method of embedding a scheme with a weight into the general difference scheme with aftereffect [10] we prove the following theorem.

Theorem 3. If, together with the stability condition $s \geq 1/2$, one applies a piecewise-linear interpolation, then method (4) converges, and the convergence order is $h^2 + \Delta$.

4. Numerical experiments. Consider the test equation

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} = \cos x \cos t - u(x, t - \pi/2)$$

with initial and boundary conditions

$$u(x, t) = \sin x \cos t, \quad 0 \leq x \leq \pi, \quad -\pi/2 \leq t \leq 0, \\ u(0, t) = 0, \quad 0 \leq t \leq \pi.$$

The exact solution of this boundary value problem is $u(x, t) = \sin x \cos t$.

The table below contains deviations $\mathbf{diff} = \max_{i,j} |u_j^i - u(x_i, t_j)|$ of the approximate solution calculated by method (4) with $s = 0.8$ from the exact one.

Table. The maximum of the absolute value of the difference between the exact and approximate solutions at mesh points

No.	1	2	3	4	5	6	7	8
h	$\pi/10$	$\pi/20$	$\pi/20$	$\pi/40$	$\pi/10$	$\pi/20$	$\pi/40$	$\pi/80$
Δ	$\pi/20$	$\pi/20$	$\pi/40$	$\pi/40$	$\pi/400$	$\pi/400$	$\pi/400$	$\pi/400$
\mathbf{diff}	0.0782	0.0971	0.0469	0.0515	0.0421	0.0144	0.0049	0.0051

In experiments Nos. 5–7 the error related to the time discretization is small in comparison with the error related to the coordinate discretization; the analysis of the error behavior reveals the square convergence with respect to x , i.e., when the step becomes half as much, the error becomes less than half as much.

The analysis of the data in the table shows that only the correlated decrease of steps yields the decrease of error. Thus, in experiments Nos. 7–8 the halving of h does not cause the corresponding decrease of error, because the total error is mostly induced by the time discretization.

By Theorem 2 for $s = 0.8$ scheme (7) is stable with any correlation of steps; however, due to the incorrectness of the numerical differentiation, the decrease of h makes the approximation of $\partial u / \partial x$ in (4) more sensitive to the computer rounding error, which leads to the increase of the error. The decrease of Δ correlated with h is a peculiar regularizer which prevents errors from growing and accumulating. Experiments Nos. 1–4 illustrate this fact.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant No. 13-01-00089) and the Analytical Departmental Target-Oriented Program 1.994.2011.

REFERENCES

1. J. Wu, *Theory and Applications of Partial Functional Differential Equations* (Springer-Verlag, New York, 1996).
2. B. Zubik-Kowal, "Delay Partial Differential Equations," http://www.scholarpedia.org/article/Delay_partial_differential_equations
3. I. B. Petrov and A. I. Lobanov, *Lectures in Computational Mathematics* (Internet-University of Information Technologies, Moscow, 2006) [in Russian].
4. L. S. Volkanin, "Numerical Solution of an Advection Equation with Heredity", in *Proceedings of Conference 'Theory of Control and Mathematical Modeling', Izhevsk, May 15–18, 2012* (IzhGTU, Izhevsk, 2012), pp. 12–14.

5. D. A. Karpenko and S. I. Solodushkin, "Numerical Solution of an Advection Equation with Aftereffect", in *International 44th All-Russian Youth School-Conference 'Modern Problems of Mathematics', Ekaterinburg, January 27 – February 2, 2013* (IMM UrO RAN, Ekaterinburg, 2013), pp. 397–400.
6. M. Gyllenberg and J. A. M. Heijmans, "An Abstract Delay Differential Equation Modelling Size Dependent Cell Growth and Division," *SIAM J. Math. Anal.* **18**, 74–88 (1987).
7. A. O. Rey and M. C. Mackey, "Multistability and Boundary Layer Development in a Transport Equation with Delayed Arguments," *Canad. Appl. Math. Quart.* **1**, 61–81 (1993).
8. A. V. Kim and V. G. Pimenov, *i-Smooth Analysis and Numerical Solution Methods for Functional-Differential Equations* (Regular and Chaotic Dynamic, Moscow–Izhevsk, 2004) [in Russian].
9. A. A. Samarskii, *Theory of Difference Schemes* (Nauka, Moscow, 1989) [in Russian].
10. V. G. Pimenov and A. B. Lozhnikov, "Difference Schemes for the Numerical Solution of the Heat Conduction Equation with Aftereffect," *Trudy Inst. Matem. i Mekhan. UrO RAN* **17** (1), 178–189 (2011).

Translated by O. V. Pinyagina